文章编号：1003-2843(2002)02-0139-05

MATLAB在数据包络分析中的应用

彭育威[1]，徐小湛[2]，吴守宪[1]
(1.西南民族学院计算机科学与技术学院，成都 610041；2.四川大学数学学院，成都 610064)

摘　要：用数学软件MATLAB编写了方便、适用的DEA应用程序，较好地解决了DEA计算量大的问题. 建立的程序为DEA理论研究和实际应用提供了方便、有效的计算工具.

关键词：数据包络分析(DEA)；线性规划；MATLAB
中图分类号：O221, O245　　　　文献标识码：A

1　DEA模型简介

数据包络分析，简称DEA (Data Envelopment Analysis)，是以相对效率概念为基础，根据多指标投入(输入)和多指标产出(输出)，对同类型的部门或单位(称为决策单元(DMU))进行相对有效性或效益评价的一种方法[1,2].

DEA是由Charnes等人于1978年提出的[3]. 该方法最初主要用于对一些非盈利部门(如教育、卫生、政府机构)的运转的有效性的评价；后来，DEA被用于更广泛的领域(如金融、经济、项目评估等等).

一个部门的运转往往需要多项投入，也会有多项产出. 例如，对大学的一个系的投入包括：教师、教师的工资、办公经费、文献资料费等等；而这个系的产出包括：培养的本科生和研究生、发表的论文、完成的科研项目等等. DEA可以对若干个同类型的这种部门或单位(它们有相同的目标和任务、有相同的输入和输出指标、有相同的外部环境)进行相对有效性的评价.

设有n个决策单元DMU_i (1● i● n). 每一个单元DMU_i有m项输入$x_{1i}, x_{2i}, \cdots, x_{mi}$ 和s项输出$y_{1i}, y_{2i}, \cdots, y_{s}$ (其中$x_{ji}, y_{ji} > 0$). 则有以下输入—输出矩阵：

	DMU_1	$\cdots$	DMU_i	$\cdots$	DMU_n
输入 1	x_{11}	$\cdots$	x_{1i}	$\cdots$	x_{1n}
输入 2	x_{21}	$\cdots$	x_{2i}	$\cdots$	x_{2n}
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
输入 m	x_{m1}	$\cdots$	x_{mi}	$\cdots$	x_{mn}
输出 1	y_{11}	$\cdots$	y_{1i}	$\cdots$	y_{1n}
输出 2	y_{21}	$\cdots$	y_{2i}	$\cdots$	y_{2n}
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
输出 s	y_{s1}	$\cdots$	y_{si}	$\cdots$	y_{sn}

将DMU_i 的输入和输出记为向量形式：

$$x_i = (x_{1i}, x_{2i}, \cdots, x_{mi})^T, \qquad y_i = (y_{1i}, y_{2i}, \cdots, y_{si})^T$$

则以上矩阵可简记为：

	DMU_1	$\cdots$	DMU_i	$\cdots$	DMU_n
输入	x_1	$\cdots$	x_i	$\cdots$	x_n
输出	y_1	$\cdots$	y_i	$\cdots$	y_n

记　$X = [x_1 \ \ x_2 \cdots x_n]$，　　$Y = [y_1 \ \ y_2 \cdots y_n]$

并称X为多指标输入矩阵，Y为多指标输出矩阵.

设　$v = (v_1, v_2, \cdots, v_m)^T$ 和 $u = (u_1, u_2, \cdots, u_s)^T$

收稿日期：2002-03-04
作者简介：彭育威(1946-)，男，西南民族学院计算机科学与技术学院教授.

分别是输入和输出的权向量，则DMU$_i$的总输入I_i和总输出O_i分别为：

$I_i = v_1x_{1i} + v_2x_{2i} + \cdots + v_mx_{mi} = x_i^T v$ 和 $O_i = u_1y_{1i} + u_2y_{2i} + \cdots + u_sy_{si} = y_i^T u$

显然，总输入I_i越小，总输出O_i越大，则DMU$_i$的效率越高．为此，DEA用总输出与总输入之比的大小来衡量DMU$_i$的有效性．令

$$E_{ii} = \frac{O_i}{I_i} = \frac{y_i^T u}{x_i^T v}$$

E_{ii}称为DMU$_i$的效率评价指数．在上式中，权向量u和v都是待定的，它们的每一个分量都是非负的(记作$u \bullet 0$, $v \bullet 0$)．对每一个DMU$_i$，我们求使E_{ii}达到最大值的权向量．因此，得到DEA的C^{2}R模型($\bar{P}$)：对每一个DMU$_i$，解以下极大化问题：

$$\begin{cases} \max \dfrac{y_i^T u}{x_i^T v} = E_{ii} \\ \text{s.t. } \dfrac{y_j^T u}{x_j^T v} \le 1 \ (1 \bullet j \bullet n),\ u \bullet 0, v \bullet 0 \end{cases} \qquad (\bar{P})$$

这是一个分式规划问题．若令

$$t = \frac{1}{x_i^T v}, \quad \bullet = tv, \quad \bullet = tu$$

则($\bar{P}$)可化为等价的线性规划问题：

$$\begin{cases} \max y_i^T \bullet = E_{ii} \\ \text{s.t. } y_j^T \bullet \bullet x_j^T \bullet \ (1 \bullet j \bullet n), \quad x_i^T \bullet = 1, \bullet \bullet 0, \bullet \bullet 0 \end{cases} \qquad (P)$$

线性规划(P)的解$\bullet_i$*和$\bullet_i$*称为DMU$_i$的最佳权向量，它们是使DMU$_i$的效率值E_{ii}达到最大值的权向量．注意：作为线性规划的解，$\bullet_i$*和$\bullet_i$*不是唯一的．

定义[2] (1)若线性规划(P)的解$\bullet_i$*，$\bullet_i$*满足：$E_{ii} = y_i^T \bullet_i^* = 1$，则称DMU$_i$为弱DEA有效(C^{2}R)的；(2)若线性规划(P)的解中存在解$\bullet_i^* > 0$，$\bullet_i^* > 0$并且$E_{ii} = y_i^T \bullet_i^* = 1$，则称DMU$_i$为DEA有效(C^{2}R)的．

为了便于检验DEA的有效性，一般考虑(P)的对偶模型的等式形式(带有松弛变量且具有非阿基米德无穷小ε)：

$$\begin{cases} \min(\theta - \varepsilon(e_1^T s^- + e_2^T s^+)) \\ \text{s.t. } \sum_{j=1}^{n} \lambda_j x_j + s^- = \theta x_i, \ \sum_{j=1}^{n} \lambda_j y_j - s^+ = y_i \\ \bullet \bullet 0, \quad s^- \bullet 0, \quad s^+ \bullet 0 \end{cases} \qquad (D_\varepsilon)$$

其中，$s^- = (s_1^-, s_2^-, \cdots, s_m^-)$是$m$项输入的松弛变量；$s^+ = (s_1^+, s_2^+, \cdots, s_s^+)$是$s$项输出的松弛变量；$\bullet = (\bullet_1, \bullet_2, \cdots, \bullet_n)$是$n$个DMU的组合系数；$e_1^T = (1,1,\cdots,1)_{1\times m}$，$e_2^T = (1,1,\cdots,1)_{1\times s}$；$\varepsilon$是一个很小的正数(一般取$\varepsilon$=10^{-6})．

定理[2] 设线性规划(D_ε)的最优解为$\bullet$*，s^{*-}，s^{*+}，$\bullet$*，则

(1)若$\bullet^*=1$，则DMU$_i$为弱DEA有效(C^{2}R)的；

(2)若$\bullet^*=1$且$s^{*-}=0$，$s^{*+}=0$，则DMU$_i$为DEA有效(C^{2}R)的．

2 MATLAB程序

由上一节知，要计算一个DMU$_i$的相对效率值并讨论其(弱)有效性，须解一个线性规划；若要计算所有DMU,

($1\bullet\ i\bullet\ n$)的相对效率值，则须解n个线性规划，其计算量比较大，一般须利用计算机进行计算．我们利用数学软件MATLAB编写了解模型(P)和(D_ε)的程序，比较方便地解决了DEA的计算量大和计算复杂的问题．

MATLAB是由Mathworks公司用C语言编写的著名的工程数学应用软件．它自1984年推向市场以来，历经十几年的发展和竞争，现已成为国际认可的最优化的科技应用软件．目前，MATLAB已经成为世界上诸多科技领域的基本应用软件．在国内、外的很多高等院校和科研机构，MATLAB已经十分普及．熟练地运用MATLAB已成为高校师生及科研人员的基本技能[4]．

MATLAB强大的矩阵运算能力和方便、直观的编程功能是我们选择它作为编写DEA应用程序的原因．诚然，LINDO或LINGO是解线性规划问题的专业软件，但它们缺乏方便的编程功能和矩阵输入功能，在解一系列线性规划时，它们不如MATLAB方便．此外，它们的普及程度远不如MATLAB．因此，我们认为MATLAB是编写DEA应用程序的最佳软件之一．

MATLAB所解的线性规划的标准形式是极小化问题：

$$\begin{cases}\min f^{*}w \\ \text{s.t. } A*w\bullet b,\ Aeq*w=beq,\ LB\bullet w\bullet UB\end{cases} \tag{1}$$

其中，w是变量，f是目标函数的系数向量，A是不等式约束的系数矩阵，Aeq是等式约束的系数矩阵，LB和UB分别是变量的下界和上界．

MATLAB解线性规划(1)的语句为：

w = LINPROG(f, A, b, Aeq, beq, LB, UB)

如果要解极大化问题max f*w，只须解极小化问题min (■f)*w．

下面，我们给出模型(P)和(D_ε)的MATLAB程序．

程序I (模型(P)的MATLAB程序)

```
clear
X=[   …     ];                %用户输入多指标输入矩阵X
Y=[   …     ];                %用户输入多指标输出矩阵Y
n=size(X',1); m=size(X,1); s=size(Y,1);
A=[-X'    Y'];
b=zeros(n,1);
LB=zeros(m+s,1); UB=[];
for i=1:n;
  f=[zeros(1,m) -Y(:,i)'];
  Aeq=[X(:,i)' zeros(1,s)]; beq=1;
  w(:,i)=LINPROG(f,A,b,Aeq,beq,LB,UB);      %解线性规划，得DMUi的最佳权向量wi
    E(i, i)=Y(:,i)'*w(m+1:m+s,i);           %求出DMUi的相对效率值Eii
end
w                             %输出最佳权向量
E                             %输出相对效率值Eii
omega=w(1:m,:)                %输出投入权向量•
mu=w(m+1:m+s,:)               %输出产出权向量•
```

程序II (模型(D_ε)的MATLAB程序)

```
clear
X=[   …     ];                %用户输入多指标输入矩阵X
Y=[   …     ];                %用户输入多指标输出矩阵Y
n=size(X',1); m=size(X,1); s=size(Y,1);
```

```
epsilon=10^-10;                %定义非阿基米德无穷小ε=10^-10
f=[zeros(1,n) -epsilon*ones(1,m+s) 1];
A=zeros(1,n+m+s+1); b=0;
LB=zeros(n+m+s+1,1); UB=[];
LB(n+m+s+1)=-Inf;
for i=1:n;
      Aeq=[X   eye(m)       zeros(m,s)     -X(:,i)
           Y   zeros(s,m)   -eye(s)        zeros(s,1)];
      beq=[zeros(m,1)
           Y(:,i)];
      w(:,i)= LINPROG (f,A,b,Aeq,beq,LB,UB);      %解线性规划，得DMUi的最佳权向量wi
end
w                                   %输出最佳权向量
lambda=w(1:n,:)                     %输出 • *
s_minus=w(n+1:n+m,:)                %输出s*-
s_plus=w(n+m+1:n+m+s,:)             %输出s*+
theta=w(n+m+s+1,:)                  %输出 • *
```

以上两个程序十分便于使用．用户只须输入多指标输入矩阵X和输出矩阵Y，即可得到所需的结果．

3 程序的应用

设有某大学的同类型的五个系 DMU_i (1• i• 5)在一学年内的投入和产出的数据如下：

		DMU_1	DMU_2	DMU_3	DMU_4	DMU_5
投入	教职工(人)	60	70	85	106	35
	教职工工资(万元)	156	200	157	263	105
	运转经费(万元)	50	180	100	86	30
产出	毕业的本科生(人)	80	60	90	96	30
	毕业的研究生(人)	12	13	20	17	8
	发表的论文(篇)	27	25	15	28	3
	完成的科研项目(项)	4	2	5	5	1

其中，运转经费指一学年内维持该系正常运转的各种费用，如行政办公费、图书资料费、差旅费等等．

由程序I，得到各系的相对效率值：

$E_{11} = 1.0000$　　$E_{22} = 0.8982$　　$E_{33} = 1.0000$　　$E_{44} = 0.8206$　　$E_{55} = 1.0000$

以及各项投入和产出的权向量：

	DMU_1	DMU_2	DMU_3	DMU_4	DMU_5
•	0.0003	0.0143	0.0001	0.0000	0.0019
	0.0002	0.0000	0.0063	0.0014	0.0015
	0.0191	0.0000	0.0001	0.0073	0.0257
•	0.0027	0.0000	0.0007	0.0000	0.0012
	0.0116	0.0554	0.0203	0.0442	0.1177
	0.0155	0.0071	0.0079	0.0000	0.0011
	0.0563	0.0000	0.0819	0.0138	0.0186

由定义，DMU_1，DMU_3和DMU_5至少是弱有效的；DMU_2和DMU_4是非弱有效的．为了确认DMU_1，DMU_3和DMU_5的有效性并分析DMU_2和DMU_4非有效的原因，须利用模型(D_ε)．

由程序 II，得本问题的解：

	DMU_1	DMU_2	DMU_3	DMU_4	DMU_5
● *	1.0000	0.8472	0.0000	1.0964	0.0000
	0.0000	0.0000	0.0000	0.0000	0.0000
	0.0000	0.1417	1.0000	0.0536	0.0000
	0.0000	0.0000	0.0000	0.0000	0.0000
	0.0000	0.0000	0.0000	0.3464	1.0000
s^{*-}	0.0000	0.0000	0.0000	4.5215	0.0000
	0.0000	25.2345	0.0000	0.0000	0.0000
	0.0000	105.1508	0.0000	0.0000	0.0000
s^{*+}	0.0000	20.5278	0.0000	6.9272	0.0000
	0.0000	0.0000	0.0000	0.0000	0.0000
	0.0000	0.0000	0.0000	3.4454	0.0000
	0.0000	2.0972	0.0000	0.0000	0.0000
● *	1.0000	0.8982	1.0000	0.8206	1.0000

由以上解可看出：DMU_1，DMU_3和DMU_5的解中● *=1且松弛变量$s^{*-}=0$，$s^{*+}=0$，故由定理知，这几个系是相对有效的．DMU_2和DMU_4的非有效性也可以在以上解中看得一清二楚．以DMU_2为例，根据有效性的经济意义[2]，在不减少各项输出的前提下，构造一个新的DMU_2：

$$DMU_2 = 0.8472^* DMU_1 + 0.1417^* DMU_3$$
$$= [\underbrace{62.8750,\quad 154.4083,\quad 56.5278}_{\text{投入}},\quad \underbrace{80.5278,\quad 13.0000,\quad 25.0000,\quad 4.0972}_{\text{产出}}]^T$$

可使DMU_2的投入按比例减少到原投入的0.8982(=● $_2$*)倍，并且(由非零的松弛变量可知)还可以进一步减少教职工工资25.2345万元、减少运转费用105.1508万元、多培养本科生20人、多完成2项科研项目．对DMU_4的非有效性可作类似的经济解释．

4 结束语

本文利用数学软件MATLAB编写了便于使用的DEA的计算程序，使DEA计算量大和计算复杂的问题得到较好的解决．本文只对DEA的C^2R模型进行了讨论．对于DEA的另一个重要模型—C^2GS^2模型，只须在模型(D_ε)中增加约束条件$\sum_{j=1}^{n}\lambda_j = 1$，程序II作相应的修改即可．本文的MATLAB程序为DEA的理论研究和实际应用提供了方便、快捷的计算工具．

参考文献：

[1] 魏权龄．评价相对有效性的DEA方法[M]．北京：中国人民大学出版社，1988.

[2] 盛昭瀚．DEA理论、方法与应用[M]．北京：科学出版社，1996.

[3] Charnes A, Cooper W W, Rhodes E. Measuring the efficiency of decision making units [J]. Eur J Opl Res, 1978, 2(6): 429-444.

[4] 许波，刘征．MatLab工程数学应用[M]．北京：清华大学出版社，2000.

MATLAB programs for DEA

PENG Yu-wei [1], XU Xiao-zhan [2], WU Shou-xian [1]

(1.College of Computer Science and Technology, Southwest University for Nationalities, Chengdu, Sichuan 610041;

2.College of Mathematics, Sichuan University, Chengdu, Sichuan 610064)

Abstract: DEA models are programmed with MATLAB. These programs offer convenient and efficient tools for DEA theories and applications.

Key words: data envelopment analysis (DEA); linear programming; MATLAB